Multilevel multivariate modelling of legislative count data, with a hidden Markov chain

Francesco Lagona, Antonello Maruotti, Fabio Padovano

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2012)
178, *Part* 3, *pp.* 705–723

Multilevel multivariate modelling of legislative count data, with a hidden Markov chain

Francesco Lagona,

Università di Roma Tre, Italy

Antonello Maruotti

University of Southampton, UK, and Università di Roma Tre, Italy

and Fabio Padovano

Université de Rennes 1, France, and Università di Roma Tre, Italy

[Received December 2012. Final revision July 2014]

Summary. The production of legislative acts is affected by multiple sources of latent heterogeneity, due to multilevel and multivariate unobserved factors that operate in conjunction with observed covariates at all the levels of the data hierarchy. We account for these factors by estimating a multilevel Poisson regression model for repeated measurements of bivariate counts of executive and ordinary legislative acts, enacted under multiple Italian governments, nested within legislatures. The model integrates discrete bivariate random effects at the legislature level and Markovian sequences of discrete bivariate random effects at the government level. It can be estimated by a computationally feasible expectation–maximization algorithm. It naturally extends a traditional Poisson regression model to allow for multiple outcomes, longitudinal dependence and multilevel data hierarchy. The model is exploited to detect multiple cycles of legislative supply that arise at multiple timescales in a case-study of Italian legislative production.

Keywords: Hidden Markov models; Longitudinal data; Multilevel models; Multivariate count data; Political legislation cycles

1. Introduction

Electoral democracies go through recurrent periods of more and less intense legislative activity, which are known as political legislation cycles. These cycles have recently been stylized by a number of studies in economics and public choice: Lagona and Padovano (2008) found evidence of pre-electoral cycles in Italian legislatures; Brechler and Gersl (2014) reached similar conclusions for the Czech Republic during the post-Communist period; evidence of political legislation cycles has also been found in France (Padovano and Gavoille, 2013) as well as in the European Parliament (Kovats, 2009).

In this literature, legislative supply is measured by univariate time series of total counts, starting at the beginning of a legislature, which are usually obtained by considering all the laws that have been enacted in each month of the legislature. The political legislation cycle is then defined as the common pattern that is shared by the legislatures under scrutiny. Under

Address for correspondence: Antonello Maruotti, Southampton Statistical Sciences Research Institute, University of Southampton, Highfield, Southampton, SO17 1BJ, UK.
E-mail: A.Maruotti@soton.ac.uk

this setting, estimating the legislation cycle is a problem of cycle extraction from panels of repeated counts: a classical problem of longitudinal data analysis. Temporal regularities of legislative supply can therefore be, at least in principle, efficiently detected, provided that the data correlation structure has been properly accounted for.

In this work, we extend this branch of the literature by looking at the legislation process as a multivariate process that evolves within a hierarchy of nested timescales. A multivariate view of the legislation process is needed because the legislative activity includes acts of various types, depending on the institutional procedures of each country. The heterogeneity of legislative acts has so far been ignored by the literature, which focuses on univariate trajectories of total legislation, which is often obtained by merging the approval of the various types of legislative instruments into a single count. This approach is reasonable under the assumption that legislators view different legislative acts as exchangeable tools to enact a given political decision. As pointed out by recent researches on political legislation theory (Padovano and Petrarca, 2013), legislators do discriminate over time between different legislative tools. A multivariate approach to legislation data analysis allows an examination of the joint distribution of multiple trajectories of legislative supply, one for each legislative act, and testing whether the hypothesis of discriminatory strategies is consistent with the data.

Moreover, during a legislature, many political events influence the legislation process. These events segment the legislatures in multiple periods and generate perturbations that take the form of small-scale cycles, nested within large-scale cycles that span through an entire legislature. Empirical evidence of nested legislation cycles has been found in the French system, where Parliamentary legislatures are interrupted by Presidential elections (Padovano and Gavoille, 2013), as well as in the European Parliamentary system, where legislatures are segmented by the reallocation of the agenda power (Kovats, 2009). Small-scale cycles seem to arise also in Italy, where legislative activity is frequently interrupted by government 'crises', i.e. changes of government due to changes in the supporting majority, without that the legislature is necessarily dissolved and new elections are called. Because of nested cycles, the trajectories of legislative activity often take complex shapes of difficult interpretation, which complicate the extraction of cycles at the legislature level. In this work, we allow for multiple timescales, facilitating the extraction of large-scale effects net of small-scale perturbations. We further exploit the additional information provided by the segmentation of a legislature, which would be ignored if a single timescale was considered.

This novel perspective on the legislative process has been made possible by a recently assembled data set on Italian legislative activity. This study involves 52 bivariate time series of the monthly counts of two different types of legislative acts, namely executive decrees and ordinary laws, which have been enacted under the 52 governments that segment the first 15 legislatures of the Italian Republic.

We examine these data by estimating a hierarchical Poisson regression model with a large-scale and a small-scale latent component. The large-scale component is specified by two correlated linear trends, one for each type of legislative act, with random coefficients that vary across legislatures. The small-scale component is instead specified as a Markov process that generates two correlated sequences of random effects, one for each type of legislative act, during each government tenure. The model hence integrates time constant correlated random effects and time varying correlated random effects within a multilevel framework, where government tenures are nested within Parliamentary legislatures.

Multilevel random effects are often used in the analysis of non-normal hierarchical data (Goldstein, 2003), as they provide a flexible strategy to account for complex correlation structures in the analysis of longitudinal data; see, for example, Skrondal and Rabe-Hesketh (2004), Steele (2008) and Muthén and Asparouhov (2009). Nevertheless, only few studies have dealt with multilevel analysis of multivariate responses (Goldstein *et al.*, 2009; Goldstein and Kounali, 2009). Correlated random-effects models are the most versatile among those that are considered in multivariate longitudinal data analysis (Verbeke *et al.*, 2014). Their potential drawback is the computational burden that is required in the estimation of the parameters, which could be made prohibitively complex as the number of random effects increases and/or as the time dependence is accounted for. Alternative solutions have been proposed by Altman (2007), Maruotti and Rydén (2009), Bartolucci and Farcomeni (2009), Maruotti (2011) and Maruotti and Rocci (2012). Severe computational issues may arise in multilevel extensions of correlated random-effects models that involve time varying random effects, like the model that we introduce for legislative data. These difficulties can be overcome by assuming a discrete distribution with a finite number of mass points for the random terms. This involves an additional step in the model selection, because the support of these distributions is not known in advance; it must be selected by evaluating the goodness of fit that different supports obtain. There are, however, several advantages that compensate such complication. First, the random-effect model reduces to a finite mixture model with a computationally tractable likelihood function. Second, the possibly inappropriate and unverifiable parametric assumptions about the distribution of the random effects are avoided. Third, the outcomes are clustered in a finite number of latent classes that can be interpreted as typical regimes of the process under examination.

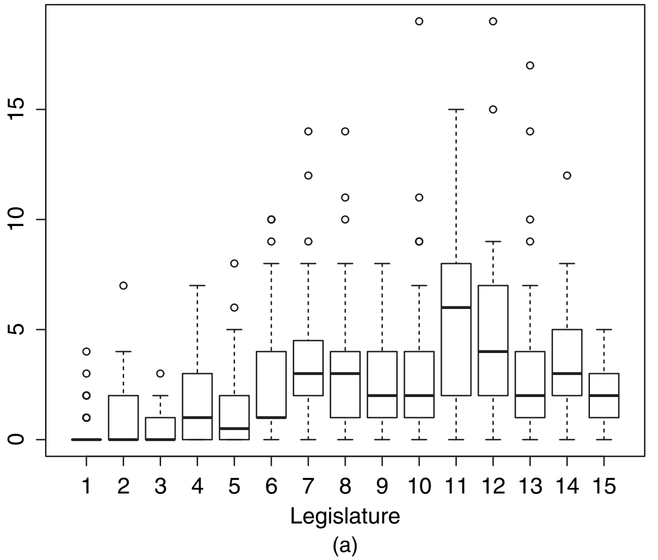
The rest of the paper is organized as follows. Section 2 reviews the specific features of the Italian legislative data that motivated the choice of a multilevel approach. The hierarchical Poisson model is illustrated in Section 3, whereas Section 4 is devoted to computational details regarding the inferential and the model selection procedures. The results are discussed in Section 5. Finally, Section 6 sums up the relevant points of discussion which emerged in the analysis.

2. Legislative data

The case-study that is discussed in this paper includes 721 months of legislative activity in Italy, observed from May 1948, when the Constitution of the Italian Republic was enacted, to May 2008, when the XVth legislature ended. This period involves 52 time series of legislative production, which begin with the appointment of a new government and end when the government resigns. Among these resignations, only 15 took place in correspondence to the end of a legislature and can therefore be considered as 'natural' events. The remaining 37 events are due to a government crisis, which eventually results in the appointment of a new government without the dissolution of the Parliament. Although the Constitution sets the normal duration of a legislature in Italy to 5 years, only five legislatures reached this deadline. The end of the remaining 10 legislatures was decided by the President of the Republic, who dissolved the Parliament before the natural deadline. Anticipated elections occur in correspondence to serious Parliamentary crises that make it impossible for the President to appoint a new government with a stable supporting majority.

We examine the two most important components of the Italian legislative activity, namely the number of executive decrees and ordinary laws that have been enacted during each month of the period under scrutiny. Decrees and laws constitute almost two-thirds of the total legislative acts in the Italian system. The main difference is that laws require the approval of the majority of the Parliament to come into effect, whereas decrees are chiefly administrative acts that usually require just the approval of the government and the ratification by the President of the Republic.

Fig. 1 shows the distribution of ordinary laws and decrees within each legislature, whereas Figs 2 and 3 display the resulting bivariate time series, clustered within the 15 legislatures under



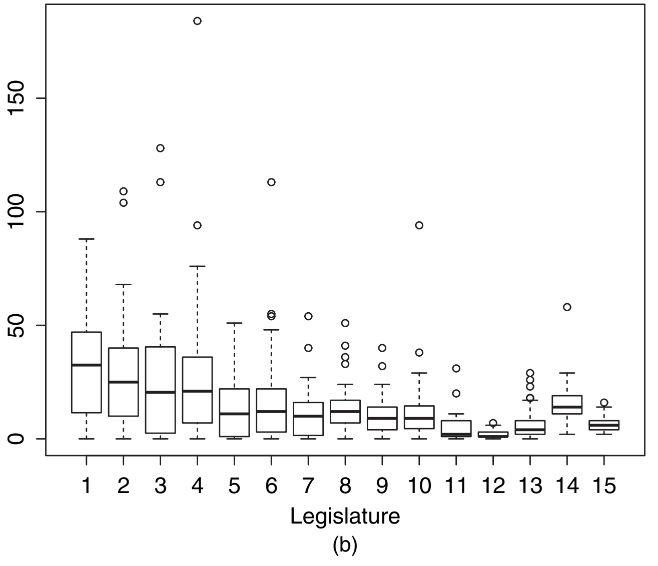


Fig. 1. Distribution of (a) executive decrees and (b) ordinary laws during the 15 legislatures of the Italian Republic

scrutiny. Figs 1–3 show also the trends estimated by fitting a Poisson regression model with the time from the beginning of a legislature as a covariate. They portray a complex picture, revealing some features of the legislation process that have been ignored by the literature so far.

First, there is a general reduction over time in the production of ordinary laws and a general increase of executive decrees (Fig. 1). This is partly explained by the need for ordinary laws in the very first years of the Italian Republic. In addition, these combined trends can be also a sign of the progressive deterioration of the Italian political system, which increasingly led legislators to use an executive decree as the standard tool to supply legislation to special interest groups, i.e. legislation that is not supported by a large consensus, thus avoiding any Parliamentary debate.

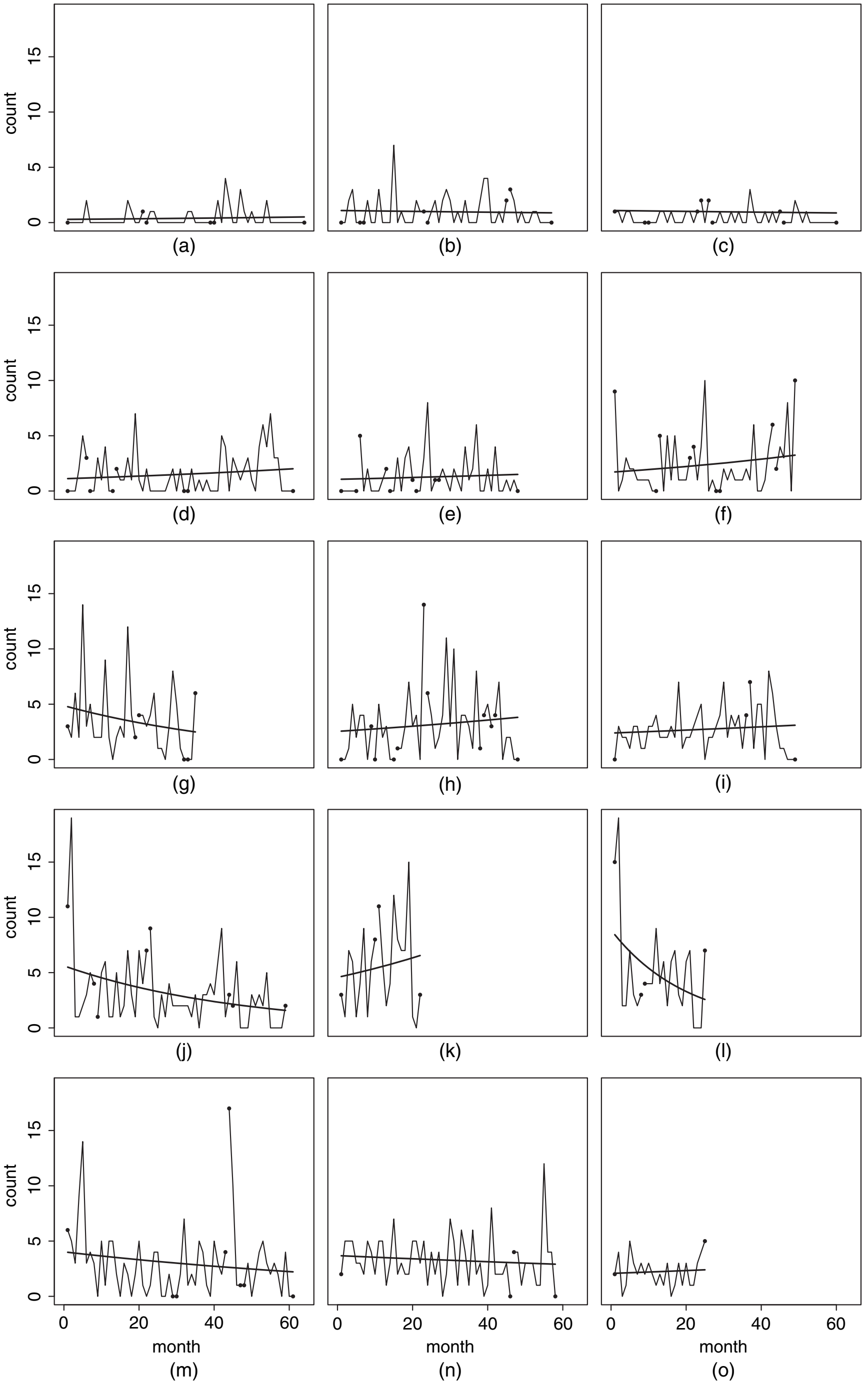


Fig. 2. Monthly counts of decrees, approved in Italy under 52 governments, during 15 legislatures (the *x*-axis indicates the time since the beginning of the legislature; ●, appointment and resignation of each government): (a) legislature 1; (b) legislature 2; (c) legislature 3; (d) legislature 4; (e) legislature 5; (f) legislature 6; (g) legislature 7; (h) legislature 8; (i) legislature 9; (j) legislature 10; (k) legislature 11; (l) legislature 12; (m) legislature 13; (n) legislature 14; (o) legislature 15

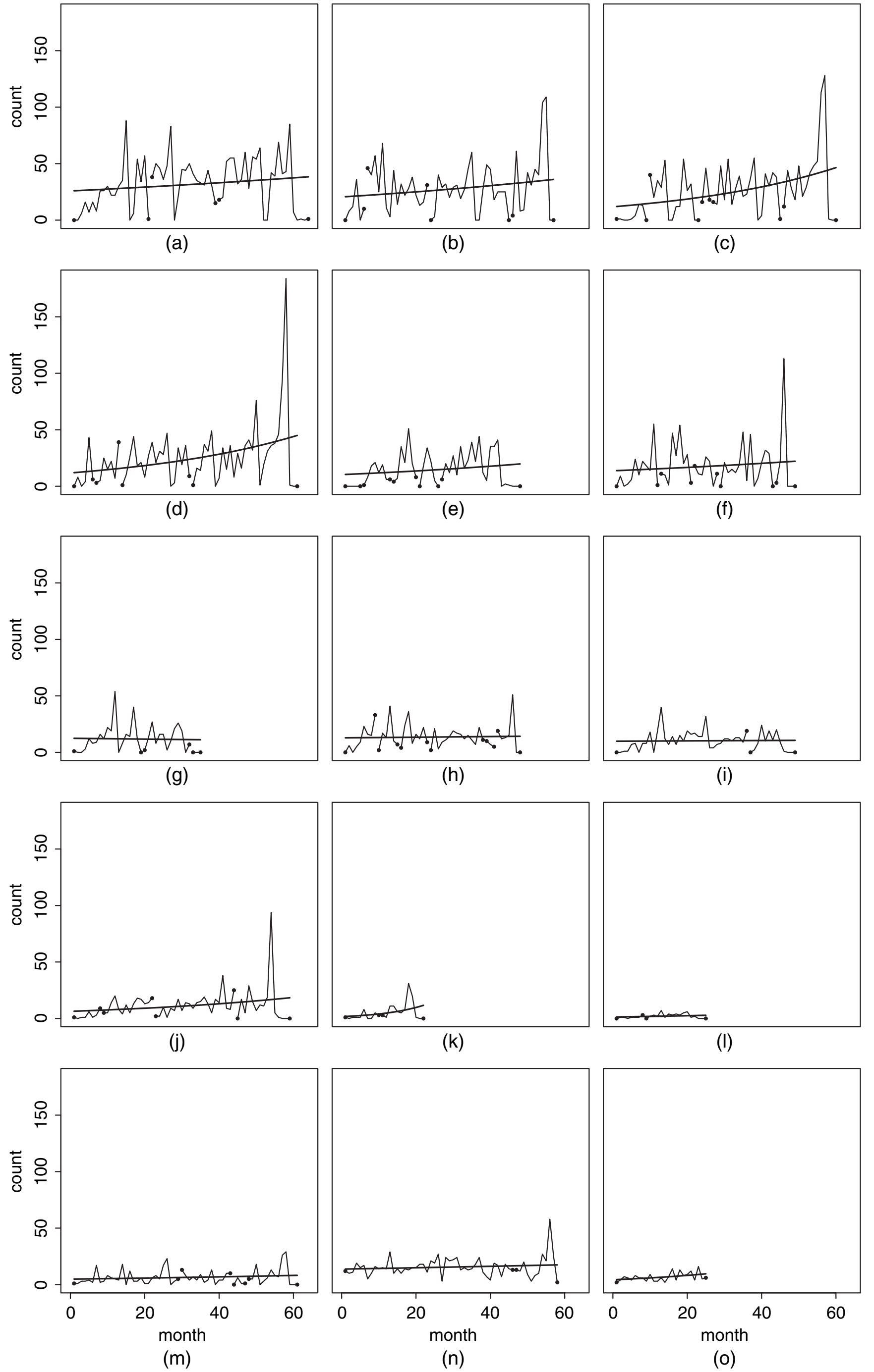


Fig. 3. Monthly counts of laws, approved in Italy under 52 governments, during 15 legislatures (the *x*-axis indicates the time since the beginning of the legislature; ●, appointment and resignation of each government): (a) legislature 1; (b) legislature 2; (c) legislature 3; (d) legislature 4; (e) legislature 5; (f) legislature 6; (g) legislature 7; (h) legislature 8; (i) legislature 9; (j) legislature 10; (k) legislature 11; (l) legislature 12; (m) legislature 13; (n) legislature 14; (o) legislature 15

Second, these trends seem almost reversed at the legislature level. Most legislatures feature an increasing production of ordinary laws as the time of the next elections approaches. The supply of decrees, in most cases, either decreases or remains constant over time. This phenomenon is, in part, consistent with the economic theory of legislation, which assumes opposite cycles of executive and ordinary legislation at the legislature level (Padovano and Petrarca, 2013). Laws are typically subject to Parliamentary debates and usually concern matters of general interest. Decrees, in contrast, go through more concealed decision-making processes and usually deal with more specific matters. The theory holds that lobbies and voters demand legislation in exchange for respectively resources and votes. Legislators supply legislation in return for resources from lobbies and votes from voters. Legislators tend to satisfy lobbies' interests at the beginning of the legislature, when re-election concerns are discounted away and resources must be gathered. Yet, as the end of the legislature approaches, time discounting raises the value of votes relative to the marginal utility of the already gathered resources and makes it increasingly optimal for legislators to satisfy voters' demands, by means of broad-based legislation. As a result, the running of time since the beginning of the legislature progressively encourages the supply of ordinary laws and, simultaneously, discourages the supply of decrees (Padovano and Petrarca, 2013).

Finally, average legislative activity seems to vary across governments, net of long-term and legislature-specific trends. This is due to the political costs that every government must sustain to implement conversely a given legislative act. These costs vary according to the competence of the members of the government, and with the size and homogeneity of the supporting majority. Inexperienced governments may find it difficult to get a law approved through a Parliamentary debate. They therefore prefer decrees as a tool to enact policy decisions. Larger government majorities find it generally easier to pass any type of decision, because of the greater Parliamentary support that they enjoy. This scale effect should be particularly evident in the case of ordinary laws, which require a Parliamentary debate.

Furthermore, government coalitions receive support at a price that increases with the degree of fragmentation of the majority. More fragmented majorities are more likely to dissolve in Parliament and therefore tend to use decrees more often than their more homogeneous counterparts, and vice versa. We have summarized these conditioning factors by means of three covariates (Table 1). The majority size MS is the minimum of the percentage of the Parliamentary seats held by each government coalition in the Chamber of Deputies and in the Senate. We consider both chambers because Italy has a system of perfect bicameralism, where to become a law each bill needs to be approved in the same reading by the two branches of the Parliament. The homogeneity index H is the degree of homogeneity of the government coalition, weighted by the fragmentation of the opposition. This index is given by $H = \mathrm{HM}(1 - \mathrm{HO})$, where HM and HO are the squared relative frequencies of the number of the overall Parliamentary seats (Chamber of Deputies plus Senate) that are held by the government and the opposition coalition respectively. Finally, the index E proxies the experience matured by the members of the government; it is computed as the average number of years in which government Ministers have served as Members of Parliament before.

It should be remarked that all factors introduced so far in the analysis vary across governments and legislatures. As such they cannot explain the fluctuations of legislative activity that we observe during the activity of a single government. During each government tenure, the observed time series appear characterized by alternating periods of intense and limited legislative activity. More interestingly, peaks of ordinary legislation often correspond to periods of moderate production of decrees; vice versa, peaks of numbers of decrees are often combined with a low supply of ordinary laws. The covariates considered may capture this heterogeneity only partially. Indeed, unobserved time varying factors may generate deviations of the outcomes from

Table 1. Italian legislative activity: summary statistics of the outcomes and the covariates

Variable	*Mean*	*Standard deviation*	*%*
Executive decrees	2.290	2.769	
Ordinary laws	16.782	19.163	
Coalition homogeneity H	0.271	0.177	
Government experience E	3.377	1.043	
Majority size MS	0.502	0.210	
Summer S			0.165

the specific trend. According to political legislation theory, these deviations can be explained in terms of different degrees of visibility of the two types of legislative acts. Laws are more 'visible' than decrees because, being subject to Parliamentary debates and concerning (usually) matters of general interest, they allow voters to become easily informed about them. Decrees, in contrast, deal with more specific matters; as such, their approval is generally known only to a more restricted set of directly interested agents, typically organized in 'lobbies'. Unorganized voters usually become informed about legislators' activities in the proximity of major political events, such as forthcoming elections or government crises. In contrast, lobbies, being characterized by lower information costs, remain always vigilant about politicians' activities. As a result, the events that draw voters' attention generate incentives to discriminate between laws and decrees and to prefer the former to the latter. This implies that, during a government tenure, the normal legislative activity can be perturbed by peaks of decrees, followed by periods of more intense production of ordinary laws.

Finally we also control for summer months, when Parliamentary recess forces the legislative activity to minimal levels. Accordingly, we have included a dummy variable S in the analysis, which equals 1 during the months of July and August and 0 otherwise (Table 1).

3. Multilevel model for longitudinal bivariate counts

The data that are considered in this paper are in the form of panels of government activity, clustered within legislatures. A clear hierarchy can be easily detected in the data structure and different sources of heterogeneity might arise at different levels of the hierarchy. Indeed, our data on legislative activity represent a special case of clustered panels, because panels are associated with the legislative activity of a specific government, and governments are sequentially appointed within a legislature. Hence panels are sequentially ordered within groups. This allows the definition of two timescales: one indicating the position of each outcome within the panel; the other indicating the position of each outcome within the cluster. Under this setting, sources of heterogeneity can be distinguished according to these two timescales, by decomposing the fluctuations of the data in small-scale and large-scale components.

More precisely, our data include n bivariate time series, which begin with the appointment of a new government and end with its resignation. These time series are clustered within L groups, each associated with a legislature. Legislatures are of different sizes and government resignations are unequally spaced during a legislature. Accordingly, we write n_l to indicate the size of the lth legislature, i.e. the number of governments appointed during this legislature, $\Sigma_{l=1}^{L} n_l = n$, whereas T_i indicates the duration (in months) of the ith government, $i = 1, \ldots, n_l$, $l = 1, \ldots, L$.

Under this setting, the data can be arranged as a multilevel array of bivariate outcomes, say

$$\mathbf{y}_{til} = (y_{til}^{(1)}, y_{til}^{(2)}), \qquad t = 1, \ldots, T_i, \qquad i = 1, \ldots, n_l, \qquad l = 1, \ldots, L,$$

where $y_{til}^{(1)}$ and $y_{til}^{(2)}$ indicate the number of decrees and of ordinary laws respectively, approved during the tth month since the appointment of government i, during legislature l. Time t indicates the position of the outcome within the time series. Each outcome is further associated with the time τ_{til} since the beginning of the lth legislature, indicating the position of the outcome within the legislature. Finally, each outcome is equipped with a row vector $\mathbf{x}_{til}^{\mathrm{T}}$ of time varying covariates. In our analysis, this vector includes the time since the enactment of the Italian Constitution, the government-specific values of the indices H, M and E and, finally, the time varying dummy variable S.

We assume that, conditionally on covariates, the legislative outcomes are generated by combining a latent small-scale process (at the government level) and a latent large-scale component (at the legislature level).

The latent small-scale process is specified by a stochastic process $(\mathbf{U}_t, t = 1, 2, \ldots)$ with bivariate components, say $\mathbf{U}_t = (U_t^{(1)}, U_t^{(2)})$, and finite dimensional distributions $p(\mathbf{u})$. Precisely, we associate each government i with two correlated sequences of random effects, one for each type of legislative act, say $\mathbf{u}_i = (\mathbf{u}_i^{(1)}, \mathbf{u}_i^{(2)})$, where

$$\mathbf{u}_i^{(1)} = (\mathbf{u}_{1i}^{(1)}, \ldots, \mathbf{u}_{ti}^{(1)}, \ldots, \mathbf{u}_{T_i i}^{(1)})$$

(decrees) and

$$\mathbf{u}_i^{(2)} = (\mathbf{u}_{1i}^{(2)}, \ldots, \mathbf{u}_{ti}^{(2)}, \ldots, \mathbf{u}_{T_i i}^{(2)})$$

(laws). These time varying random effects are drawn from the T_i-dimensional distribution $p(\mathbf{u}_{T_i})$ of the bivariate process $\mathbf{U}_t$ and represent small-scale fluctuations of legislative activity.

The latent large-scale component is instead specified by two correlated trends, with random coefficients that vary across legislatures, independently of the small-scale process $\mathbf{U}_t$. Precisely, we associate each legislature l with two correlated trends, one for each legislative act, say

$$\begin{aligned} y_\tau^{(1)} &= v_{0l}^{(1)} + v_{1l}^{(1)}\tau, \\ y_\tau^{(2)} &= v_{0l}^{(2)} + v_{1l}^{(2)}\tau, \end{aligned}$$

where τ is the time since the beginning of legislature l and the four coefficients $\mathbf{v}_l = (v_{0l}^{(1)}, v_{1l}^{(1)}, v_{0l}^{(2)}, v_{1l}^{(2)})$ are drawn from a four-dimensional distribution $p(\mathbf{v})$.

In summary, the large-scale component is specified by an array $\mathbf{v} = (\mathbf{v}_1 \ldots \mathbf{v}_L)$ of L four-dimensional, legislature-specific random effects. The small-scale component is specified by an array $\mathbf{u} = (\mathbf{u}_1 \ldots \mathbf{u}_n)$ of government-specific trajectories of time varying, bivariate random effects.

We further assume that the responses are conditionally independent given the large-scale and the small-scale random effects, or, in other words, that the conditional distribution of all the responses $\mathbf{y} = (y_{til}, t = 1, \ldots, T_i, i = 1, \ldots, n_l, l = 1, \ldots, L)$ is a product of univariate conditional distributions, say

$$p(\mathbf{y}|\mathbf{v}, \mathbf{u}) = \prod_{j=1}^{2} \prod_{l=1}^{L} \prod_{i=1}^{n_l} \prod_{t=1}^{T_i} p(y_{til}^{(j)}|\mathbf{v}, \mathbf{u}).$$

In particular, we assume that each univariate conditional distribution depends on the random effects only through $v_{0l}^{(j)}, v_{1l}^{(j)}$ and $u_{ti}^{(j)}$, namely

$$p(y_{til}^{(j)}|\mathbf{v},\mathbf{u}) = p(y_{til}^{(j)}|v_{0l}^{(j)}, v_{1l}^{(j)}, u_{ti}^{(j)}).$$

Under these assumptions, the joint distribution of the observed and the unobserved quantities, say

$$p(\mathbf{y},\mathbf{v},\mathbf{u}) = \prod_{l=1}^{L} p(\mathbf{v}_l) \prod_{i=1}^{n_l} p(\mathbf{u}_i) \prod_{t=1}^{T_i} \prod_{j=1}^{2} p(y_{til}^{(j)}|v_{0l}^{(j)}, v_{1l}^{(j)}, u_{ti}^{(j)}), \tag{1}$$

is fully specified by defining

(a) the univariate conditional distributions of each outcome given covariates and random effects,
(b) the multivariate distribution of the legislature-specific random effects and
(c) the multivariate distribution of the government-specific trajectories of random effects.

We specify the univariate response distributions by augmenting the linear predictor of a Poisson generalized linear model with the random effects introduced above. Precisely, we assume that the responses $y_{til}^{(j)}$ are drawn from a Poisson distribution with mean $\lambda_{til}^{(j)}$, linked to fixed and random effects through the canonical logarithm transformation, namely

$$\begin{gathered} p(y_{til}^{(j)}|v_{0l}^{(j)}, v_{1l}^{(j)}, u_{ti}^{(j)}) \sim \mathrm{Poi}(\lambda_{til}^{(j)}), \\ \log(\lambda_{til}^{(j)}) = \mathbf{x}_{til}^{\mathrm{T}}\boldsymbol{\beta}_j + v_{0l}^{(j)} + v_{1l}^{(j)}\tau_{til} + u_{ti}^{(j)} \end{gathered} \tag{2}$$

where $\boldsymbol{\beta}_j$ is a vector of (outcome-specific) regression coefficients and $t=1,\ldots,T_i$, $i=1,\ldots,n_l$, $l=1,\ldots,L$, $j=1,2$. The above specification is identifiable under standard identifiability constraints, namely $v_{01}^{(j)} = u_{1i}^{(j)} = 0$, $i=1,\ldots,n_l$, $l=1,\ldots,L$, $j=1,2$.

To specify the distributions of the random terms, we take a non-parametric approach and assume that the random effects are drawn from discrete distributions with a finite number of mass points. At the legislature level, we therefore assume that the random effects $\mathbf{v}_l$ are independently drawn for each legislature from a discrete distribution, say

$$\mathbf{v}_l \sim p_M(\mathbf{v},\boldsymbol{\pi}), \tag{3}$$

where $p_M(\mathbf{v},\boldsymbol{\pi})$ depends on a vector of M support points $\mathbf{v}=(\mathbf{v}_1\ldots\mathbf{v}_M)$ with mass probabilities $\boldsymbol{\pi}=(\pi_1\ldots\pi_M)$. Each support point includes four co-ordinates, namely $\mathbf{v}_m=(v_{0m}^{(1)}, v_{0m}^{(2)}, v_{1m}^{(1)}, v_{1m}^{(2)})$, $m=1,\ldots,M$. Under this setting, at the legislature level, each support point is associated with a latent class that indicates a pair of correlated trends: one for each type of legislative act. At the government level, we assume that the process $(\mathbf{U}_t, t\geqslant 1)$ is a Markov chain with K bivariate states $\mathbf{u}=(\mathbf{u}_1\ldots\mathbf{u}_K)$, $\mathbf{u}_k=(u_k^{(1)}, u_k^{(2)})$. Each state of the chain is therefore associated with a pair of random effects, one for each legislative act, that indicate a specific latent regime of legislative activity. Because of the Markov property, the chain distribution is fully known up to an initial probability distribution $\boldsymbol{\delta}=(\delta_1\ldots\delta_K)$, $\delta_k=P(\mathbf{U}_1=\mathbf{u}_k)$, and a $K\times K$ matrix $\mathbf{Q}$ of transition probabilities $q_{kh}=\Pr(\mathbf{U}_t=\mathbf{u}_h|\mathbf{U}_{t-1}=\mathbf{u}_k), t>1, h,k=1,2,\ldots,K, \Sigma_{h=1}^{K}q_{kh}=1$. Accordingly, for each government i, we assume that the sequences $\mathbf{u}_i=(\mathbf{u}_{ti}, t=1,\ldots,T_{il})$ are independently drawn from a joint discrete distribution, say

$$\mathbf{u}_i \sim p_K(\mathbf{u},\boldsymbol{\delta},\mathbf{Q}), \tag{4}$$

where $p_K(\mathbf{u},\boldsymbol{\delta},\mathbf{Q})$ is the joint distribution of $(\mathbf{U}_1\ldots\mathbf{U}_t\ldots\mathbf{U}_{T_i})$.

This model includes some popular approaches to longitudinal data analysis as particular cases. For example, when the number of the latent states K is equal to 1, the model reduces to a (bivariate) latent class growth model (Nagin and Land, 1993), in which bivariate panels are clustered within M latent classes, each associated with a pair of correlated parametric trends. On the other side, when the distribution $p(\mathbf{v})$ concentrates the whole probability mass on the origin, $p(\mathbf{v}=\mathbf{0})=1$, the model reduces to a (bivariate) hidden Markov model (Bartolucci and Farcomeni, 2009; Bartolucci *et al.*, 2013), in which bivariate panels are segmented by the sequences of a Markov chain with K states. As an intermediate case, if $p(\mathbf{v})$ concentrates the whole probability mass on the intercept, $\Sigma_{m=1}^{M} p(v_{0m}^{(1)}, v_{0m}^{(2)}, 0, 0)=1$, the model reduces to a (bivariate) mixed hidden Markov model (Maruotti, 2011), where the outcomes in a given panel are segmented by sequences of K latent states and, simultaneously, share a common intercept, chosen among the M mass points of distribution $p(\mathbf{v})$.

4. Likelihood inference

The multilevel Poisson model that was illustrated in Section 3 depends on a vector $\boldsymbol{\theta}$ of parameters that includes three components:

(a) the fixed effects $\boldsymbol{\beta}_j$, $j=1,2$,
(b) the multivariate support points $(v_{0m}^{(1)}, v_{0m}^{(2)}, v_{1m}^{(1)}, v_{1m}^{(2)})$ and the related mass probabilities $\boldsymbol{\pi}$ of the random-effects distribution at the legislature level and
(c) the bivariate support points $(u_k^{(1)}, u_k^{(2)})$ of the time varying random effects, with the initial probabilities, collected in $\boldsymbol{\delta}$, and the transition probabilities, collected in $\mathbf{Q}$, of the hidden chain.

The maximum likelihood estimate of $\boldsymbol{\theta}$ is the maximum point of the marginal likelihood function

$$L(\boldsymbol{\theta})=\prod_{l=1}^{L} \sum_{\mathbf{v}} p_M(\mathbf{v};\boldsymbol{\pi}) \prod_{i=1}^{n_l} \sum_{\mathbf{u}} p_K(\mathbf{u};\boldsymbol{\delta},\mathbf{Q})\, p(\mathbf{y}_{il}|\mathbf{u},\mathbf{v},\boldsymbol{\beta}) \tag{5}$$

obtained by integrating the joint distribution (1) with respect to the random effects, where

$$p(\mathbf{y}_{il}|\mathbf{u},\mathbf{v},\boldsymbol{\beta})=\prod_{t=1}^{T_i} \prod_{j=1}^{2} p(y_{til}^{(j)}|v_{0l}^{(j)}, v_{1l}^{(j)}, u_{ti}^{(j)}, \boldsymbol{\beta}_j).$$

Expression (5) can be efficiently computed by an extension of the forward recursion, which is well known in the hidden Markov model literature (Baum *et al.*, 1970).

To maximize the likelihood $L(\boldsymbol{\theta})$, we implement a version of the expectation–maximization (EM) algorithm (Baum *et al.*, 1970), which is facilitated by the independence assumption between the latent states of the Markov chain and the legislature-specific random effects. The algorithm is based on a complete-data likelihood function, which is obtained by considering the logarithm of the joint distribution of both the observed and the unobserved quantities (1). It can be conveniently written in terms of dummy variables that indicate class membership. Accordingly, let η_{lm} denote a dummy variable that is equal to 1 if the lth legislature is in latent group m. Similarly, let ξ_{tik} denote a dummy variable that is equal to 1 if the ith government is in latent state k at time t and $\zeta_{tikk'}=\xi_{t-1,ik}\xi_{tik'}$ be a dummy variable that is equal to 1 if there is a transition from latent state k to latent state k' at time t. Under this setting, the complete-data log-likelihood of the model proposed is given by

$$l_{\mathrm{c}}(\boldsymbol{\theta})=\sum_{l=1}^{L} \sum_{m=1}^{M} \eta_{lm} \log(\pi_m) \tag{6}$$

$$+\sum_{i=1}^{n} \sum_{k=1}^{K} \xi_{1ik} \log(\delta_k) \tag{7}$$

$$+\sum_{i=1}^{n}\sum_{k=1}^{K}\sum_{k'=1}^{K}\sum_{t=2}^{T_i}\zeta_{tikk'}\log(q_{kk'}) \tag{8}$$

$$+\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k=1}^{K}\sum_{i=1}^{n_l}\sum_{t=1}^{T_i}\eta_{lm}\xi_{tik}\sum_{j=1}^{J}\log\{p(y_{til}^{(j)}|\mathbf{x}_{til},\mathbf{v}_m,\mathbf{u}_k)\}. \tag{9}$$

The EM algorithm alternates two steps until convergence: first, the conditional expected value of the complete-data log-likelihood is computed (the E-step); then, this expectation is maximized with respect to $\boldsymbol{\theta}$. The EM algorithm is guaranteed to converge to a local maximum of the likelihood. To increase the chances of reaching the global maximum, we use a short-run strategy (which was described in the case of Gaussian mixtures by Biernacki *et al.* (2003)) and run the algorithm from various starting values, stopping the algorithm without reaching the full convergence. We then select the output which maximizes the log-likelihood and then use such a value to initialize longer runs of the EM algorithm. Finally, we use the value of the likelihood at convergence to perform model selection according to the Bayesian information criterion (BIC) (Schwarz, 1978).

In the E-step, the conditional expected value of terms (6)–(9) is simply computed by a plug-in of the expected values of η_{lm}, ξ_{tik} and $\zeta_{tikk'}$ given the observed data and the current value of the parameters. Such quantities can be computed by means of an appropriate forward–backward recursion adapted from the mixed hidden Markov model literature (Maruotti, 2011). In the M-step, the conditional expected value of terms (6)–(9) is maximized by separately maximizing its terms. It is straightforward to verify that explicit solutions are available for the latent parameters. In particular, at iteration $r+1$, we have

$$\pi_m^{(r+1)} = \frac{\sum_{l=1} \mathbb{E}(\eta_{lm}|\mathbf{y},\boldsymbol{\theta}^{(r)})}{L},$$

$$\delta_k^{(r+1)} = \frac{\sum_{i=1}^{n} \mathbb{E}(\xi_{1ik}|\mathbf{y},\boldsymbol{\theta}^{(r)})}{n},$$

$$q_{kk'}^{(r+1)} = \frac{\sum_{i=1}^{n}\sum_{t=2}^{T_i} \mathbb{E}(\zeta_{tikk'}|\mathbf{y},\boldsymbol{\theta}^{(r)})}{\sum_{i=1}^{n}\sum_{t=2}^{T_i} \mathbb{E}(\xi_{t-1,ik}|\mathbf{y},\boldsymbol{\theta}^{(r)})}.$$

As for the regression parameters and the latent locations, we can use Newton–Raphson algorithms, similar to that used for standard generalized linear models. Formally, the update of these parameters is obtained by maximizing

$$\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{k=1}^{K}\sum_{i=1}^{n_l}\sum_{t=1}^{T_i}\mathbb{E}(\eta_{lm}|\mathbf{y},\boldsymbol{\theta}^{(r)})\,\mathbb{E}(\xi_{tik}|\mathbf{y},\boldsymbol{\theta}^{(r)})\sum_{j=1}^{J}\log\{p(y_{til}^{(j)}|\mathbf{x}_{til},\mathbf{v}_m,\mathbf{u}_k)\}. \tag{10}$$

The maximization of this expression is not complex. Indeed, formula (10) corresponds to weighted sums of $M\times K$ likelihood equations for standard generalized linear models with weights given by the posterior probabilities $\mathbb{E}(\eta_{lm}|\mathbf{y},\boldsymbol{\theta}^{(r)})$ and $\mathbb{E}(\xi_{tik}|\mathbf{y},\boldsymbol{\theta}^{(r)})$. The procedure outlined above does not provide standard errors for the estimated parameters. We therefore consider a parametric bootstrap approach, i.e. we have refitted the model to 250 bootstrap samples, which are simulated from the estimated model parameters. Other approaches can be pursued (see for example Bartolucci and Farcomeni (2014)).

Table 2. Model selection

Government effect	*Time varying heterogeneity*	*Legislature effect*	*Log-likelihood*	*BIC*
No	No	No	−6378.74	12859.31
2 classes	No	No	−6052.76	12217.39
3 classes	No	No	−5952.63	12036.87
4 classes	No	No	−5928.68	12008.71
5 classes	No	No	−5917.01	12005.12
No	No	2 classes	−6208.93	12608.70
No	No	3 classes	−6128.34	12414.62
No	No	4 classes	−6056.78	12304.39
2 classes	No	2 classes	−6103.82	12352.42
2 classes	No	3 classes	−6060.72	12299.11
2 classes	No	4 classes	−6044.61	12299.80
3 classes	No	2 classes	−5991.59	12147.70
3 classes	No	3 classes	−5939.82	12077.05
3 classes	No	4 classes	−5865.47	11961.25
4 classes	No	2 classes	−5941.86	12067.97
4 classes	No	3 classes	−5855.71	11928.58
4 classes	No	4 classes	−5769.09	11788.25
5 classes	No	2 classes	−5868.18	11940.36
5 classes	No	3 classes	−5843.06	11923.01
5 classes	No	4 classes	−5750.37	11770.54
2 states	Yes	No	−4929.91	9984.85
3 states	Yes	No	−4235.50	8642.10
4 states	Yes	No	−4061.30	8352.92
5 states	Yes	No	−3915.79	8134.29
6 states	Yes	No	−3894.43	8177.12
2 states	Yes	2 classes	−4815.38	9788.70
2 states	Yes	3 classes	−4748.94	9688.72
2 states	Yes	4 classes	−4727.55	9678.84
3 states	Yes	2 classes	−4180.16	8564.32
3 states	Yes	3 classes	−4113.71	8464.32
3 states	Yes	4 classes	−4078.44	8426.69
4 states	Yes	2 classes	−3975.86	8214.95
4 states	Yes	3 classes	−3940.99	8178.11
4 states	Yes	4 classes	−3921.87	8172.77
5 states	Yes	2 classes	−3876.26	8088.13
5 states	Yes	3 classes	−3844.48	8057.48
5 states	Yes	4 classes	−3837.00	8075.42
6 states	Yes	2 classes	−3870.19	8161.55
6 states	Yes	3 classes	−3831.61	8117.28
6 states	Yes	4 classes	−3707.99	8082.95

5. Results

When dealing with discrete random effects, and latent class models in general, it is crucial to identify the latent structure that recovers the unobservable data structure in the most accurate and interpretable way. We use the BIC to select the optimal number of latent government-specific states and legislature-specific groups. Furthermore, the BIC is also used to highlight the usefulness and appropriateness of the approach proposed, by comparing more parsimonious multilevel and single-levels models, which ignore selected components of the model proposed. The results of the model selection procedure are displayed in Table 2. Firstly, we consider a fixed effects approach, by omitting government- and legislature-specific random effects, i.e. assuming independent observations. This leads to a misspecified model. Simply adding (time

Table 3. Multilevel model with $K = 5$ and $M = 3$: parameter estimates

		Results for decrees		*Results for laws*	
		Estimate	*Standard error*	*Estimate*	*Standard error*
Covariate effects	Majority size	0.640	0.148	0.567	0.070
	Coalition homogeneity	−0.536	0.194	0.583	0.061
	Government experience	−0.057	0.026	0.124	0.014
	Summertime	−0.187	0.072	−0.688	0.037
Long-run trend	Intercept	−0.376	0.197	2.919	0.076
	Overall time (month)	0.003	0.000	−0.002	0.000
Legislature level: latent effects	Intercept class 2	−0.896	0.146	0.476	0.054
	Intercept class 3	0.142	0.107	−0.551	0.057
	Slope class 1	0.007	0.003	0.020	0.001
	Slope class 2	0.005	0.004	0.017	0.001
	Slope class 3	−0.019	0.003	0.030	0.001
Government level: latent effects	Intercept state 2	−1.592	0.190	−5.021	0.207
	Intercept state 3	0.816	0.103	−3.320	0.134
	Intercept state 4	−0.340	0.095	−1.484	0.028
	Intercept state 5	0.174	0.089	−0.810	0.022

constant) specific effects in only one of the hierarchy levels improves the fit considerably, and the BIC is strongly in favour of a random-effects model specification that allows for dependent observations. Further improvement of the goodness of fit is provided by multilevel models that integrate random effects at the government and at the legislature level, confirming the presence of different sources of heterogeneity. Among these, the best results are obtained with the model proposed, where the random effects at the government level are dependent over time through a Markov chain.

From this analysis, a multilevel model with $K = 5$ states of time varying government-specific effects and $M = 3$ classes of legislature-specific random effects seems a good compromise between goodness of fit and parsimony. This model decomposes the heterogeneity within the legislation process into four components: first, the effects of a set of covariates, predicted by the economic theory of legislation, that operate at the single government level within each legislature; second, two long-run trends that span the 15 legislatures, one for each type of legislative instrument; third, three pairs of legislature-specific trends, which identify three different patterns of legislative activity; fourth, and finally, government-specific sequences of random effects, which capture the residual variability of the supply of legislation during the tenure of each government. Table 3 displays the estimates of these four components.

The covariate effects are consistent with the predictions of the economic theory of legislation, which was described in Section 2. The size of the Parliamentary majority acts as a scale factor, since it facilitates, other things being equal, the approval of both types of legislative instrument. As expected, this effect is much more significant for ordinary laws, which require the approval of the Parliament. The homogeneity of the government coalition, instead, has opposite effects on the supply of decrees and of laws. More homogeneous coalitions tend to pass policies through ordinary laws, as they are less fearful of an open Parliamentary debate. By the same logic, more fragmented majorities tend to resort to decrees more. The experience and the competence of government Ministers present a similar pattern of results; high rates of production of ordinary laws are associated with high levels of competence or experience, since more seasoned Ministers are likely to have more followers in the Parliament; low levels of competence or experience are instead correlated with high rates of approval of executive legislation. Finally, as expected, the summer months cut down the enactment of both decrees and laws; yet, this drop is more evident for laws than for decrees. This difference is possibly because of the Parliament recess in summer, whereas ministries remain open during this period. It could also indicate that legislators take advantage of the fact that during summer months voters' attention is lowest, which minimizes the cost of satisfying lobbies' interests by means of decrees.

The second part of Table 3 illustrates two opposite long-term trends across legislatures: an upward sloping trend for decrees and a downward sloping trend for ordinary laws. These trends well represent the progressive deterioration of the Italian political system. As time went by, Italian legislators increasingly refrained from taking decisions through an open Parliamentary debate and preferred to adopt them by means of less visible legislative acts, such as decrees.

The third part of Table 3 displays the coefficients of the $M=3$ trends of legislative activity that operate at the level of each legislature, with $\boldsymbol{\pi}=(0.53; 0.20; 0.27)$. These trends can be interpreted as components of large-scale variation, net of the long-run trends, of the fixed effects and of government-specific small-scale fluctuations. Each trend is associated with a latent class; the posterior membership probabilities $\mathbb{E}(\eta_{lm}|\mathbf{y},\boldsymbol{\theta}^{(r)})$ can be exploited to allocate each legislature to a maximum probability class. This allows associating each trend with a homogeneous group of legislatures. Specifically, the model allocates legislatures II, IV, V, VI, VII, VIII, IX and XI to group 1, legislatures I, XIV and XV are clustered within group 2, and the remaining legislatures are assigned to group 3. The distinguishing feature of group 1, the group with the largest number of legislatures, is an increasing supply of ordinary laws as the months pass and a steady supply of decrees over time. Group 2 shares similar slopes to those of group 1, but the intercepts are quantitatively different, as they identify a higher production of laws and a lower supply of decrees. Group 3, finally, is instead characterized by opposite cycles of decrees and laws.

The finding of three distinct types of legislature clearly shows that the various legislatures do not have a uniform political legislation cycle, as both the theoretical and the empirical literatures often maintain. It also points out that the opposite cycles of laws and decrees, which the economic theory of legislation predicts, take place only under special conditions. Group 3 actually includes most of the legislatures that reached their natural end of 5 years. However, nine of the 10 legislatures that ended prematurely do not show opposite cycles of laws and decrees. The small number (15) of legislatures in our sample, however, prevents us from firmly establishing that the probability of observing opposite legislation cycles increases with the duration of a legislature. Nonetheless, this result is in line with Lagona and Padovano (2008) who, in a similar sample, found that the natural end of the legislature is the necessary condition to observe a peak of production of ordinary laws immediately before the elections.

A possible interpretation of these results is that political instability, combined with uncertainty about the end of the legislature, keeps voters' attention alive throughout the entire duration of the legislature. This in turn discourages legislators from targeting special interests groups by means of less visible decrees, which therefore appear evenly spread throughout the legislature. At the same time, this persistent public attention pushes legislators to enact ordinary laws that cater for broad interests at rates which increase with the probability of an imminent political crisis.

Finally, the model clusters the monthly fluctuations of legislative activity of each government within $K=5$ latent states. Again, the maximum over k of the posterior probabilities $\mathbb{E}(\xi_{tik}|\mathbf{y},\boldsymbol{\theta}^{(r)})$ is used to allocate each month to one of these five states. Each state is associated with a pair of intercepts, which reflect the deviations of the rates of production of decrees and laws from the legislature-specific trends. Table 3 displays these deviations, taking as reference the pair $(-0.376, 2.919)$, of the legislature-specific trends in group 1. The resulting state-specific intercepts are shown in Fig. 4(a). Three typical regimes can be distinguished. State 2 reflects low levels of production of both laws and decrees. An intense supply of executive legislation, combined with a reduced production of ordinary laws, characterizes state 3. Finally, the remaining three states (1, 4 and 5) reveal the opposite behaviour, i.e. an above-trend production of laws and a below-trend production of decrees. The small-scale fluctuations of the legislative activity of the various governments can be best described as an alternation of regimes of either moderate (state 2) or intense legislative activity (the other states). This is consistent with the patterns that have already been described in Figs 2 and 3. Moreover, during the periods of intense legislative activity, the deviations of decrees and laws are negatively correlated. This suggests that legislators discriminate between laws and decrees within a timescale that is defined by the life span of a single government and not only of an entire legislature. The choice of a legislative instrument is determined by time varying latent factors that operate between the appointment and the resignation of a government, even if these events do not correspond to the beginning or the end of a legislature. Examples may be a looming government crisis or the approval of the budget bill. These factors are likely to include political events that occasionally draw the public attention and create incentives to discriminate between more or less visible legislative tools.

The remaining three panels of Fig. 4 illustrate the timing at which these discriminatory strategies take place. For this, we show the percentages of governments in a given state, computed every month since the day of their appointment. These trajectories are cut at the 13th month, which is the median survival time of the governments in the sample. Beyond this point, the small sample sizes lead to irregular trajectories with misleading shapes. In this computation, we have also merged the percentages of states 1, 4 and 5, which are associated with similar regimes of legislative activity. The trajectory of high production of ordinary legislation (states 1, 4 and 5) increases after the 10th month. This suggests that a looming government crisis draws voters' attention and anticipates the approval of highly visible legislation. As a government crisis approaches, however, governments become increasingly unable to supply legislation of any type; transitions to regimes of modest activity (state 2) appear more often. Conversely, regimes of high production of executive decrees (state 3) follow a decreasing trajectory. In the first months of activity, when the new executive enjoys a sort of 'honeymoon' period, governments tend to concentrate on satisfying lobbies' interests through acts with low visibility. After this period, governments tend to move to the opposite regime characterized by intense ordinary legislation (states 1, 4 and 5). Overall, this pattern of results shows that small-scale fluctuations of legislative activity are characterized by opposite cycles of decrees and laws. Specifically, we estimate the following initial and transition probabilities of the government level latent Markov chain:

$$\delta = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \end{pmatrix} = \begin{pmatrix} 0.04 \\ 0.22 \\ 0.34 \\ 0.28 \\ 0.12 \end{pmatrix},$$

$$Q = \begin{pmatrix} q_{11} & q_{12} & q_{13} & q_{14} & q_{15} \\ q_{21} & q_{22} & q_{23} & q_{24} & q_{25} \\ q_{31} & q_{32} & q_{33} & q_{34} & q_{35} \\ q_{41} & q_{42} & q_{43} & q_{44} & q_{45} \\ q_{51} & q_{52} & q_{53} & q_{54} & q_{55} \end{pmatrix} = \begin{pmatrix} 0.18 & 0.31 & 0.04 & 0.12 & 0.35 \\ 0.01 & 0.50 & 0.11 & 0.31 & 0.07 \\ 0.03 & 0.26 & 0.16 & 0.38 & 0.17 \\ 0.06 & 0.03 & 0.00 & 0.42 & 0.49 \\ 0.16 & 0.02 & 0.04 & 0.26 & 0.52 \end{pmatrix}.$$

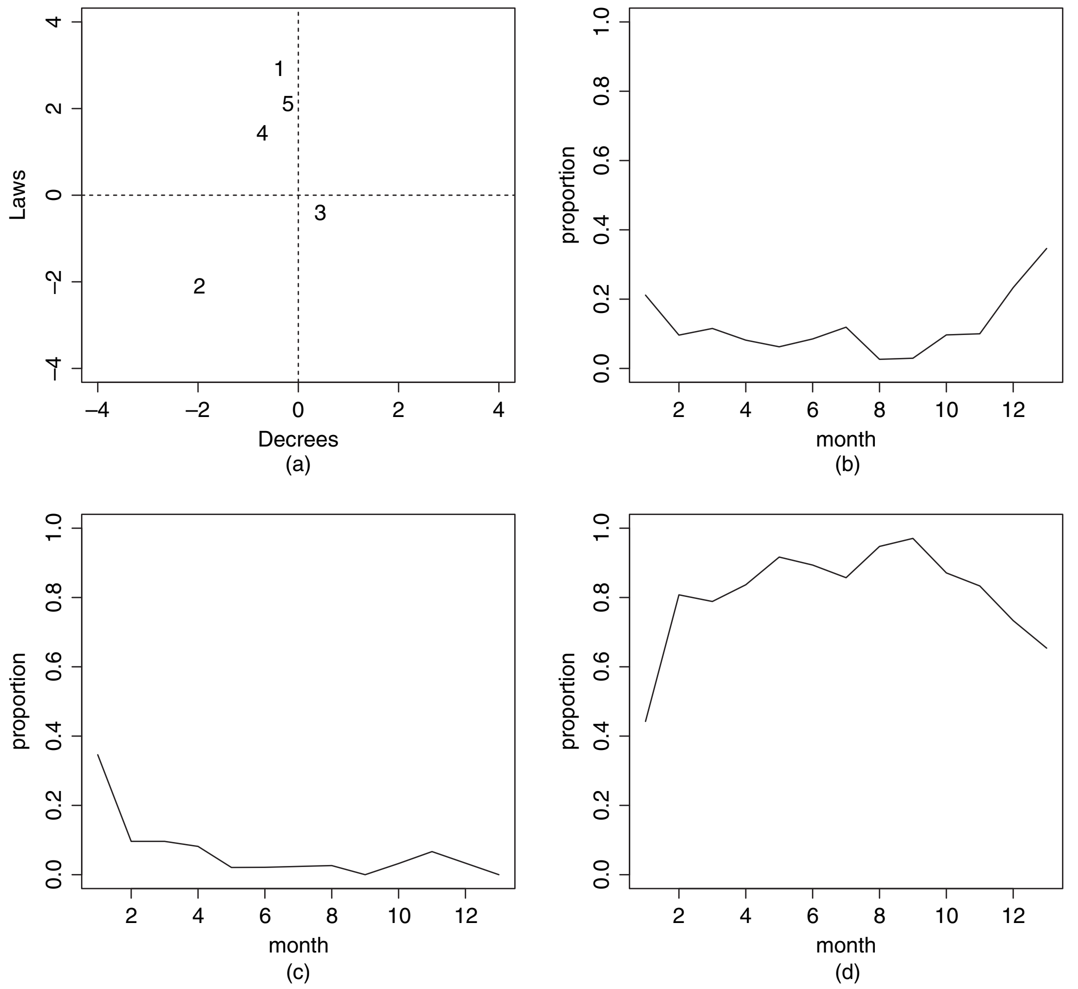


Fig. 4. (a) The five small-scale regimes of legislative activity under a government, and proportions of governments in (b) state 2, (c) state 3 and (d) states 1, 4 and 5 during the first 13 months since the appointment

In summary, this analysis sheds new light on the complex nature of the legislation process, by providing some empirical evidence that the existing literature had not yet detected. First, legislators discriminate between different legislative instruments over time. Univariate analyses of total legislation are misleading. Second, temporal regularities of multivariate legislative activity emerge net of the effects of the size and homogeneity of the majority and of the competence of the government. These regularities appear consistent with the economic theory of legislation. Third, latent factors affect the legislative process at multiple timescales, with effects of government behaviour nested within large-scale effects that characterize the legislature. Additionally, the model detects two long-term trends that are specific to each legislative instrument that endure across legislatures. Complex trajectories of legislative activity can hence be explained only through the combination of multiple components that act at different timescales.

6. Discussion

Combining different sources of latent heterogeneity is a popular strategy to model complex correlation structures. This approach is particularly convenient when multiple components of heterogeneity are suggested by the structure of the data. In legislation studies, for example, the data often take the form of panels clustered within groups. Latent heterogeneity can then be modelled by combining two components that respectively reflect the variation between panels and the variation within panels.

We consider a case-study of legislation where non-normal outcomes are bivariate, panels are temporally ordered within groups and the interest lies in the estimation of large-scale trajectories of legislative activity that arise at the group level, net of small-scale patterns within each panel. Taking advantage of the data structure, we propose a novel multilevel Poisson model that integrates a latent growth model for bivariate trajectories and a latent Markov chain for bivariate small-scale patterns. From a methodological viewpoint, the model proposed extends that proposed by Bartolucci and Farcomeni (2009), who introduced a time varying latent structure for longitudinal data (without a multilevel structure). In particular, our proposal flexibly accommodates temporal correlation within panels, heterogeneity between clusters of panels and cross-correlations between two non-normal outcomes. Despite this complexity, the model is easy to interpret, as it combines models that are broadly exploited in longitudinal data analysis. It is also relatively easy to estimate, since it is based on discrete distributions of random effects that feature a computationally tractable likelihood function.

Taking a multilevel–multivariate approach to legislation data analysis provides some empirical evidence that the existing literature has not detected so far. Temporal regularities of legislative supply emerge net of the effects of the size and homogeneity of the majority and of the competence of the government. These regularities appear as a result of the time varying strategic interactions between legislators, lobbies and unorganized voters. Distinct patterns of legislative activity occur in combination with the relevant political events that segment the activity of a political system. We provided such evidence by examining the Italian case. Further research is needed to compare these findings with those that may emerge from data about other Western democracies.

A major issue in the approach considered is related to model selection. In the current paper we consider the BIC to select the *best* model. Nevertheless, as is clear from the results discussed, the BIC tends to overestimate the order of the model, which is defined as the number of distinct latent classes at the different levels of the hierarchy. Indeed, three out of five identified latent states at the government level capture similar behaviours and two out of three latent classes at the legislature level identify similar trends. This is because the BIC selects the latent structure that is needed to provide a good approximation to the probability distribution, rather than the number of homogeneous behaviours as such. This is a well-known problem in multivariate clustering (see for example Baudry *et al.* (2010)) and further research is needed on model selection methods in a multivariate–multilevel context.

We assumed that the random effects at different levels are generated by independent processes. This assumption leads to a parsimonious multilevel model. An interesting extension would be to avoid such a restriction as for example in Altman (2007). Nevertheless, including random effects simultaneously in the observed and latent processes may lead to flat likelihoods, and convergence may be difficult to achieve. In addition, although appealing, this specific extension may not be necessary when the sample size is limited (as in our empirical application).

Acknowledgements

The authors thank the Associate Editor and the referees for the detailed review of our paper and the very useful comments. They have been precious for improving the quality of the paper.

The authors are grateful to Dankmar Böhning and Valentina Raponi for their comments on a draft version of this paper.

References

Altman, R. J. (2007) Mixed hidden Markov models: an extension of the hidden Markov model to the longitudinal data setting. *J. Am. Statist. Ass.*, **102**, 201–210.
Bartolucci, F. and Farcomeni, A. (2009) A multivariate extension of the dynamic logit model for longitudinal data based on a latent Markov heterogeneity structure. *J. Am. Statist. Ass.*, **104**, 816–831.
Bartolucci, F. and Farcomeni, A. (2014) Information matrix for hidden Markov models with covariates. *Statist. Comput.*, to be published, doi 10.1007/s11222-014-9450-8.
Bartolucci, F., Farcomeni, A. and Pennoni, F. (2013) *Latent Markov Model for Longitudinal Data*. Boca Raton: CRC Press.
Baudry, J. P., Raftery, A. E., Celeux, G., Lo, K. and Gottardo, R. (2010) Combining mixture components for clustering. *J. Computnl Graph. Statist.*, **19**, 332–353.
Baum, L., Petrie, T., Soules, G. and Weiss, N. (1970) A maximization technique occurring in the statistical analysis of probabilistic function of Markov chains. *Ann. Math. Statist.*, **41**, 164–171.
Biernacki, C., Celeux, G. and Govaert, G. (2003) Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computnl Statist. Data Anal.*, **41**, 561–575.
Brechler, J. and Gersl, A. (2014) Political legislation cycle in the Czech Republic. *Constitnl Polit. Econ.*, **25**, 137–153.
Goldstein, H. (2003) *Multilevel Statistical Models*. Chichester: Wiley.
Goldstein, H., Carpenter, J., Kenward, M. and Levin, K. (2009) Multilevel models with multivariate mixed response types. *Statist. Modllng*, **9**, 173–197.
Goldstein, H. and Kounali, D. (2009) Multilevel multivariate modelling of childhood growth, numbers of growth measurements and adult characteristics. *J. R. Statist. Soc.* A, **172**, 599–613.
Kovats, L. (2009) Do elections set the pace?: a quantitative assessment of the timing of European legisation. *J. Eur. Publ. Poly*, **16**, 239–255.
Lagona, F. and Padovano, F. (2008) The political legislation cycle. *Publ. Choice*, **134**, 201–229.
Maruotti, A. (2011) Mixed hidden Markov models for longitudinal data: an overview. *Int. Statist. Rev.*, **79**, 427–454.
Maruotti, A. and Rocci, R. (2012) A mixed non-homogeneous hidden Markov model for categorical data, with application to alcohol consumption. *Statist. Med.*, **31**, 871–886.
Maruotti, A. and Rydén, T. (2009) A semiparametric approach to hidden Markov models under longitudinal observations. *Statist. Comput.*, **19**, 381–393.
Muthén, B. and Asparouhov, T. (2009) Multilevel regression mixture analysis. *J. R. Statist. Soc.* A, **172**, 639–657.
Nagin, D. S. and Land, K. C. (1993) Age, criminal careers and population heterogeneity: specification and estimation of a nonparametric mixed Poisson model. *Criminology*, **31**, 327–362.
Padovano, F. and Gavoille, N. (2013) The dual political legislation cycle in France. *Working Paper 2013-02-ccr*. Centre Condorcet for Political Economy, University of Rennes 1, Rennes.
Padovano, F. and Petrarca, I. (2013) A theory of cyclical production of laws and decrees. *Constitnl Polit. Econ.*, **24**, 336–351.
Schwarz, G. (1978) Estimating the dimension of a model. *Ann. Statist.*, **6**, 461–464.
Skrondal, A. and Rabe-Hesketh, S. (2004) *Generalized Latent Variable Modelling: Multilevel, Longitudinal and Structure Equation Models*. Boca Raton: Chapman and Hall–CRC.
Steele, F. (2008) Multilevel models for longitudinal data. *J. R. Statist. Soc.* A, **171**, 5–19.
Verbeke, G., Fieuws, S., Molenberghs, G. and Davidian, M. (2014) The analysis of multivariate longitudinal data: a review. *Statist. Meth. Med. Res.*, **23**, 42–59.